Stability of graph communities across time scales

J.-C. Delvenne[1*], S.N. Yaliraki[1,2], and M. Barahona[1,3]
[1]*Institute for Mathematical Sciences,*
[2]*Department of Chemistry and* [3]*Department of Bioengineering*
Imperial College London, South Kensington Campus,
London SW7 2AZ, United Kingdom

(Dated: May 5, 2010)

Supplementary Information

I. STABILITY PROVIDES A DYNAMICAL INTERPRETATION OF OTHER HEURISTICS IN GRAPH CLUSTERING AND COMMUNITY DETECTION

The definition of stability introduced in our work establishes a unifying framework by providing a dynamical interpretation of several clustering heuristics that have been proposed in diverse areas of the literature, usually based on combinatorial or statistical arguments. In the main body of the paper we have concentrated on the explicit connections with some of the most popular measures, such as modularity. However, there are several other interesting interpretations, to which we alluded in passing in the text, and we now consider those in more detail.

Recalling the definitions in the text, the stability $r(t)$ of the graph is defined in terms of the clustered autocovariance matrix

$$R_t = H^T \left(\Pi M^t - \pi^T \pi\right) H \tag{1}$$

as follows:

$$r(t; H) = \min_{0 \leq s \leq t} \sum_{i=1}^{c} (R_s)_{ii} = \min_{0 \leq s \leq t} \text{trace}\left[R_s\right], \tag{2}$$

and the stability curve of the network is obtained by maximizing this measure over all possible partitions:

$$r(t) = \max_H r(t; H).$$

Note that the definition of stability naturally incorporates:

- the *topology of the graph*, which is encoded in the matrix $M = D^{-1}A$, where A is the adjacency matrix and $D = \text{diag}(d_j)$ with d_j the degree of node j, and also in the stationary distribution $\pi = \pi M$, which is the left eigenvector of M with unit eigenvalue;
- the *partition of the graph* through the indicator matrix H.

As explained in the text, an undirected, unweighted graph with $m = \sum_j d_j/2$ edges leads to an ergodic reversible Markov chain with $\pi_j = d_j/2m$. This implies that $\Pi = D/2m$ and $\Pi M = A/2m$.

Other measures and their dynamical interpretation in terms of Stability:

Explicit connections of stability, as a dynamical definition of the goodness of a partition, with other static (combinatorial or statistical) measures that have been proposed in the literature follow below:

1. **Stability at time zero is the Diversity Index:**

   The stability at time zero is

$$r(0; H) = \text{trace}\left[H^T (\Pi - \pi^T \pi) H\right]$$

[*] Present address: Department of Applied Mathematics, Université catholique de Louvain, 4, av. Lemaître, B-1348 Louvain-la-Neuve, Belgium

and it is easy to see that

$$\text{trace}\left[H^T \Pi H\right] = H^T D H / 2m = 1,$$

which leads to the result

$$r(0; H) = 1 - \sum_i (\pi h_i)^2 = \text{Diversity Index}.$$

Note that this quantity is essentially identical to the *diversity index*. There have been several formulations of this measure in different areas of the literature: Simpson's diversity index [1] (used in population biology), the Hirschman-Herfindahl index [2] (used in economics) or the Rényi entropy of order 2 [3] (used in information theory and statistical mechanics). These measures are all equivalent to the above definition, up to a simple transformation.

The Diversity Index can be seen as a natural and popular way to measure how heterogeneous or diverse a population is. This measure is useful to characterize the diversity of an ecological system, for instance, or the importance of monopolistic practices in a market. It is easy to see that the Diversity Index is maximal when the graph is split in N communities with one node. This index has the flavor of an entropy and it favors having as many communities as possible and tries to make them of equal size (in terms of degrees). In the case of graphs with a heterogeneous degree list, putting the nodes of highest degree in the same community decreases the Diversity Index maximally.

2. **Stability at time one is Modularity:**

   As stated in the text, it follows from our definition that

   $$r(1; H) = \text{trace}\left[H^T (\Pi M - \pi^T \pi) H\right] = \text{Modularity}$$

   is exactly *modularity*, as introduced in [4].

3. **The variation of Stability between time zero and time one is the Cut size:**

   If we subtract modularity from the diversity index, the resulting quantity:

   $$\text{Diversity Index} - \text{Modularity} = r(0; H) - r(1; H) = 1 - \text{trace}\left[H^T \Pi M H\right] = \text{Cut}$$

   has a simple interpretation as the *cut size*, i.e., the fraction of edges between communities. Indeed, it is easy to show that

   $$\text{trace}\left[H^T \Pi M H\right] = \frac{\text{trace}\left[H^T A H\right]}{2m} = \sum_i \frac{h_i^T A h_i}{2m},$$

   which is the fraction of edges inside communities.

4. **Modularity as a compound cost function of Cut and Diversity Index**:

   It follows from above that:

   $$\text{Modularity} = \text{Diversity Index} - \text{Cut}.$$

   This provides an interesting interpretation of modularity as a cost function with two competing objectives: when modularity is maximized, this is equivalent to minimizing the cut size (the negative term favors small cut sizes), while trying to maximize the Diversity Index (the first term favors a large number of equal-size communities in terms of edges), thus leading to more balanced cuts.

5. **The linearization of Stability is equivalent to the Multi-resolution Potts heuristic**:

   The Potts Hamiltonian used in [5] as a heuristic to find multi-resolution community structure in graphs can be rewritten as

   $$\mathcal{H}_\gamma = -2m \ \text{trace}\left[H^T (\Pi M - \gamma \pi^T \pi) H\right],$$

   with the tunable resolution parameter γ. It is easy to see from our results above that this Potts 'energy' can be directly interpreted in terms of a compound cost function:

   $$-\frac{\mathcal{H}_\gamma}{2m} = (1 - \gamma) + (\gamma \,\text{Diversity Index} - \text{Cut}).$$

This means that minimizing the Potts energy is equivalent to maximizing the function (γ Diversity Index $-$ Cut), i.e., it is equivalent to minimizing the Cut with a tunable constraint on the Diversity Index, with Lagrange multiplier γ, that regulates how balanced the resulting partitions are.

As mentioned in the text, our definition of stability also provides a dynamical interpretation of the Potts Hamiltonian. It is straightforward to see that a linear interpolation of the stability between its value at $t = 0$ and $t = 1$ has the following direct correspondence:

$$r_c(t) = r(0; H)\,(1-t) + r(1; H)\,t = (1-t) - t\left(\frac{1}{2m}\mathcal{H}_{1/t}\right).$$

Therefore maximizing $r_c(t)$, the (linear interpolation) of stability at time t, is equivalent to minimizing the Potts Hamiltonian $\mathcal{H}_{1/t}$. This also gives an explicit interpretation of the heuristic multi-resolution parameter γ in terms of the Markov time: $\gamma = 1/t$. It is important to remark that $r_c(t)$ can also be viewed as the linearization of the associated continuous Markov process for small times[6].

It is also interesting to remark that our linearization of the stability $r_c(t)$ can be rewritten as

$$-\frac{r_c(t)}{t} = \text{Cut} - \frac{1}{t}\,\text{Diversity Index}.$$

Since the Diversity Index can be viewed as an entropy and Cut can be considered as an energy cost, this indicates that the linearization of stability has in effect the flavor of a *free energy* to be minimized, with $1/t$ playing the role of temperature.

6. **The linearization of Stability is related to the Self-loop modified modularity of Arenas *et al*:**

The linear interpolation above can also be rewritten as

$$r_c(t) = r(0; H)\,(1-t) + r(1; H)\,t = \text{trace}\left[H^T\left(\Pi D^{-1}\left((1-t)D + tA\right) - \pi^T\pi\right)H\right],$$

which is precisely the modularity of a modified graph with adjacency matrix $\widetilde{A} = (1-t)D + tA$, where D is the diagonal matrix of degrees. This means that maximizing the linearization of stability $r_c(t)$ at time t is equivalent to maximizing the modularity of a new graph in which self-loops are added to each node j of the graph with non-uniform strength $d_j(1-t)/t$.

This is related to the scheme proposed by Arenas *et al* [7] in which they studied the modularity of $A + rI$, with a parameter $r > 0$, in order to discover smaller communities than the ones found for A. The Arenas scheme is thus tantamount to studying the modularity associated with the adjacency matrix $\frac{1-t}{t}I + A$ and it is equivalent to our linear interpolation of stability when the graph is regular.

7. **Normalized Cut is related to a version of Stability with autocorrelations:**

As stated in the main text, stability can be viewed as a sum of autocovariance functions restricted to the prescribed communities, as enforced by the indicator matrix H. For instance, if we consider a partition into two communities U_1 and U_2 and denote the corresponding indicator matrix as H_{U_1/U_2}, the stability of the partition is given by

$$r(t, H_{U_1/U_2}) = \min_{s\leq t}\left[(R_s)_{11} + (R_s)_{22}\right].$$

If we now replace the autocovariances by the corresponding autocorrelations, we get a new quantity, which we denote $\rho(t)$ instead of $r(t)$. In the case of a two-way partition as above,

$$\rho(t, H_{U_1/U_2}) = \min_{s\leq t}\left[\frac{(R_s)_{11}}{\pi(U_1) - \pi(U_1)^2} + \frac{(R_s)_{22}}{\pi(U_2) - \pi(U_2)^2}\right].$$

Here $\pi(U_1) = \sum_{j\in U_1}\pi_j$ is the total probability of U_1 and $\pi(U_2) = \sum_{j\in U_2}\pi_j$ is the total probability of U_2. It then follows that $\pi(U_1) - \pi(U_1)^2 = \pi(U_2) - \pi(U_2)^2 = \pi(U_1)\pi(U_2) = (\pi(U_1)^{-1} + \pi(U_2)^{-1})^{-1}$. We then observe that

$$\rho(t, H_{U_1/U_2}) = \frac{r(t, H_{U_1/U_2})}{(\pi(U_1)^{-1} + \pi(U_2)^{-1})^{-1}},$$

from which it is immediate to show that

$$\rho(0, H_{U_1/U_2}) - \rho(1, H_{U_1/U_2}) = \mathrm{Cut}\ \left(\pi(U_1)^{-1} + \pi(U_2)^{-1}\right) = 2\,\mathrm{NCut},$$

where NCut is the Normalized Cut size introduced in [8], as a heuristic for balanced partitions. Hence NCut is related to the difference between $t = 0$ and $t = 1$ of the version of stability with autocorrelations, much the same way as Cut can be expressed as the difference between $t = 0$ and $t = 1$ of the standard stability. We remark that the above relation does not hold for more than two communities.

8. **The asymptotic limit of Stability and its relation to the Normalized Fiedler Vector:**

Stability also has interesting connections with spectral methods. Spectral methods are classic heuristics in the clustering literature and have been widely used for decades. More recently, spectral methods have also been invoked in the optimization of modularity (see below). Our definition of stability provides a link to these heuristics through a dynamical interpretation. We expand on some of these explicit connections below.

Consider the clustered autocovariance matrix R_t that lies at the heart of our definition of stability:

$$R_t = H^T \left(\Pi M^t - \pi^T \pi\right) H \tag{3}$$

with $M = D^{-1}A$, where $A_{N\times N}$ is the adjacency matrix of the graph, and $D = \mathrm{diag}(\mathbf{d})$, where $\mathbf{d} = A\mathbf{1}$ is the vector of node degrees, $\mathbf{1}_{N\times 1}$ is the vector of ones and the total number of edges is $m = \mathbf{d^T 1}/2$. If the graph is undirected and unweighted, then $\pi_{1\times N} = \mathbf{d^T}/2m$, implying that $\Pi = D/2m$ and $\Pi M = A/2m$. Note that the particular clustering is encoded in the indicator matrix $H_{N\times c}$, where c is the number of communities in the clustering.

We now consider the spectral decomposition of the symmetrized matrix of the process above (3):

$$\mathcal{M} = D^{1/2} M D^{-1/2} = \sum_{i=1}^{N} \lambda_i \mathbf{u}_i \mathbf{u}_i^T\,,$$

where the eigenvalues λ_i are ranked in decreasing order ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$), and the corresponding eigenvectors $\mathbf{u}_i$ are orthonormal. In particular, it is easy to see that $\lambda_1 = 1$ and $\mathbf{u}_1 = (1/\sqrt{2m})\, D^{1/2}\mathbf{1}$. Note that $\mathcal{M}$ and M are related via a similarity transformation and their spectra are identical. Note also that $\mathcal{M} = I - \mathcal{L}$, where $\mathcal{L} = D^{-1/2} L D^{-1/2}$ is the normalized Laplacian and $L = D - A$ is the combinatorial Laplacian. Hence the spectra of $\mathcal{M}$ (and M) are trivially related to that of the normalized Laplacian $\mathcal{L}$.

The dynamics (3), which governs stability, can then be expressed as:

$$R_t = \frac{1}{2m} H^T D^{1/2} \left(\sum_{i=2}^{N} \lambda_i^t \mathbf{u}_i \mathbf{u}_i^T \right) D^{1/2} H\,.$$

If we assume for simplicity that R_t is monotonically decreasing, as observed in all the examples presented in this article, then it follows that $r(t;H) = \mathrm{trace}\, R_t$. (This occurs for instance if all eigenvalues $\lambda_i \geq 0$.) Furthermore, if there is no degeneracy in the top eigenvalues, $\lambda_1 > \lambda_2 > \lambda_3$ (see the discussion in the Remarks below), then the stability behaves asymptotically as:

$$r_\infty \equiv \lim_{t\to\infty} R_t \sim \frac{\lambda_2^t}{2m} \| H^T D^{1/2} \mathbf{u}_2 \|^2\,, \tag{4}$$

which is governed by $\mathbf{u}_2$, the *normalized* Fiedler eigenvector, and its corresponding eigenvalue.

We can now study what partition optimizes this asymptotic quantity by quantifying the variation induced by associating nodes into communities. First consider the 'finest' partition possible, with each node in its own community. The indicator matrix of this partition is the identity matrix, $H_f = I_{N\times N}$, and the asymptotic value of its stability follows from:

$$r_{\infty,f} \sim \frac{\lambda_2^t}{2m} \| D^{1/2} \mathbf{u}_2 \|^2 = \frac{\lambda_2^t}{2m} \sum_{i=1}^{N} d_i u_{2,i}^2.$$

If we consider the 'next finest' clustering, where we group two nodes i and j in the same community, this asymptotic value is then:

$$r_{\infty,nf} \sim \frac{\lambda_2^t}{2m} \left(\sum_{k\neq i,j} d_k u_{2,k}^2 + \left(\sqrt{d_i} u_{2,i} + \sqrt{d_j} u_{2,j} \right)^2 \right).$$

It then follows that the variation in the asymptotic stability induced by going from the 'finest' to the 'next finest' partitions, i.e., by putting nodes i and j in the same community, is:

$$r_{\infty,nf} - r_{\infty,f} \sim \frac{\lambda_2^t}{m} \sqrt{d_i d_j}\, u_{2,i}\, u_{2,j},$$

which only increases if

$$\text{sign}\left[u_{2,i}\right] = \text{sign}\left[u_{2,j}\right].$$

Therefore the partition that maximizes the asymptotic form of the stability is the one given by grouping the nodes according to the sign of the components of the *normalized* Fiedler vector $\mathbf{u}_2$, which is a recently proposed heuristic [8] for graph partitioning and hierarchical clustering (and an extension of the classic result of Fiedler's [9, 10]).

Remarks about the asymptotic behavior:

- **Stability indicates when the Normalized Fiedler bi-partition, although optimal, is not a good community:** The above result indicates that as time goes to infinity, stability will tend towards coarser partitions and in the $t \to \infty$ limit, a bi-partition according to the sign of the normalized Fiedler vector is favored. However, this does *not* imply that such a two-way partition will be considered a good community since the value of its stability when dominant could be so low as to be negligible, in which case this 2-way partition will not be selected as a good community. This point is discussed extensively in the next section.
- **Special cases when the asymptotic behavior is not dominated by the Normalized Fiedler vector:** The asymptotic dominance of the normalized Fiedler eigenvector is typical for realistic graphs, even when R_t is not monotonically decreasing. There are however two important cases when this is not true and in which stability reveals features of the structure of the network:

  Graphs with degenerate second eigenvalues: When the second eigenvalue has multiplicity two ($\lambda_2 = \lambda_3$) or higher, the asymptotic behavior is governed by the subspace spanned by the degenerate eigenvectors of the normalized Laplacian. In this case, the coarsest optimal community structure will typically comprise more than two communities. When λ_2 and λ_3 are nearly degenerate, i.e., very close to each other compared to the gap between λ_1 and λ_2, a multi-way partition will emerge as the dominant community structure for very long times, thus implying that the coarser partitions (and specifically the 2-way partition) will not appear as relevant communities. A (near) degeneracy of eigenvalues is often induced by a (near) symmetry in the graph. This point is also discussed extensively with examples in the next section.

  Bipartite and almost bipartite graphs: If a graph is bipartite, i.e., when it can be split into two groups that have edges between them but no edge inside of them, then $\lambda_1 = |\lambda_N| > \lambda_2$. An 'almost bipartite' graph is characterized by $\lambda_1 > |\lambda_N| > \lambda_2$. In this case, it follows that R_t behaves asymptotically as $\frac{\lambda_N^t}{2m}\|H^T D^{1/2}\mathbf{u}_N\|^2$, and, since $\lambda_N < 0$, it will therefore oscillate around zero at large times. Based on our definition (2), the stability $r(t, H)$ will therefore be negative at large enough times for all partitions except for the one-way partition, for which $H = \mathbf{1}$ which implies that its stability is identically zero. This means that the one-way partition would be asymptotically optimal in the case of bipartite (or almost bipartite) graphs—a reasonable outcome for such graphs where there is no clear 2-way cut to be enacted.

  It is also interesting to note that in the case of almost bipartite graphs, the asymptotic limit of stability is typically *minimized* by maximizing $\frac{\lambda_N^t}{2m}\|H^T D^{1/2}\mathbf{u}_N\|^2$, which, similarly to above, is optimized by sorting the vertices according to the sign of u_N. In the case of bipartite, this leads to a bipartition with no edge inside and all edges across (or into 'anti-communities'). So, for bipartite (and almost bipartite) graphs, stability maximization leads to the analysis of the graph in terms of 'anti-community' structure. Hence stability can be related to communities when maximized and to 'anti-communities' when minimized.

II. COMMUNITY DETECTION AND GRAPH PARTITIONING

A. Communities *vs* partitions

As explained in [11, 12], the problem of *graph partitioning* is related to, but distinct from that of *community detection.* In broad terms, the difference revolves around the fact that in graph partitioning the goal is to find the best division of a network when the number and size of the sub-graphs is imposed 'a priori', while in community detection the number and size of the sub-graphs are not imposed 'a priori' and are determined as an outcome of the algorithm itself.

A typical partitioning problem is to find a partition into two sets of vertices of equal sizes such that the number of edges between the two sets is minimal. If one were to use graph-partitioning methods as a means to finding communities, one could obtain graph partitions that are not truly relevant as communities. An example of this would be to use the classical Fiedler method to find bi-partitions of a graph, even if there is no natural two-way community for the network.

Communities should thus emanate from the properties of the network rather than from an a priori choice of the external user. In fact, communities may not exist at all. This is another difference between graph partitioning and community detection: partitioning algorithms should find an answer on every graph considered (the best possible partition according to a heuristic), while a community detection algorithm is expected to assess the quality of the answer, and conclude if needed that there is no significant community in the graph.

B. Stability as a measure of good communities disregards irrelevant partitions

One of the key properties in the definition of modularity is that it does not pre-determine the number of communities found in a graph[11]. Our definition of stability has the same property and does not impose the number of communities a priori either.

As explained in the main text, stability has the ability to compare partitions into different numbers of communities (even produced according to different heuristic methods), and determine if they become meaningful over the whole span of timescales. As a rule of thumb, one would expect 'fine' partitions to be relevant at short times and 'coarse' partitions to be relevant at long timescales.

However, it is important to remark that stability does not impose the emergence of particular partitions as communities, or indeed of any community, as a necessary outcome. In fact, the value of stability can be used to establish the relevance of partitions: partitions that are not relevant as communities will have very low values of their stability. This means that such partitions will not be chosen as meaningful communities on any timescale. It is therefore possible that no partitions are found to be good communities if the graph has no clear community structure.

In particular, the asymptotic behavior as $t \to \infty$ does *not* imply necessarily that the normalized Fiedler two-way partition will be a relevant community. As our examples below show, if the normalized Fiedler bi-partition is not a good community, its stability will decay very fast and will be always negligible for practical purposes. If there are other relevant partitions (but not two-way), this will be reflected in the existence of quasi-degeneracy in the spectrum of the normalized Laplacian. This will have the effect that λ_2 will not dominate the convergence of the dynamics until extremely long times for which the stability of the bi-partition will be negligible. Hence, we emphasize that this two-way partition is not imposed as a necessary outcome of the community detection.

As stated above, our definition of stability implies in general that finer partitions tend to be optimal at short times while coarser partitions tend to be optimal at longer times. The asymptotic behavior means that if we were to consider the top non-trivial level of a hierarchy of clusterings and we assumed it to be composed of only two communities, the two-way partition favored asymptotically would be given by normalized Fiedler. In light of our discussion above, establishing such a level on the hierarchy would mean that two-way partitioning and community detection would be indistinguishable in the limit of large times. However, stability would allow us to establish that such a partition is not relevant as a community.

We exemplify these behaviors through the following cases:

1. **Graphs with no relevant community structure:**

   If a network does not have genuine community structure, the stability of the partitions decays quickly. Hence the partitions will not be characterized as relevant communities. For example, a random graph does not have any relevant community structure (bar unavoidable fluctuations) as seen in Figure I below, where we compare the stability curves of the protein graph analyzed in the paper (2085 nodes) against the stability of a random graph of the same size and similar average degree. Our example shows that, contrary to the behavior of the protein graph, the stability of all partitions quickly converges to zero in the random graph. This is a clear indication that no significant community structure is present in the random graph, beyond any accidental structure due to unavoidable fluctuations. Similar results apply to expander graphs, which by definition are regular graphs with large spectral gap in which all partitions have a high conductance, hence implying that these graphs have no meaningful community structure.

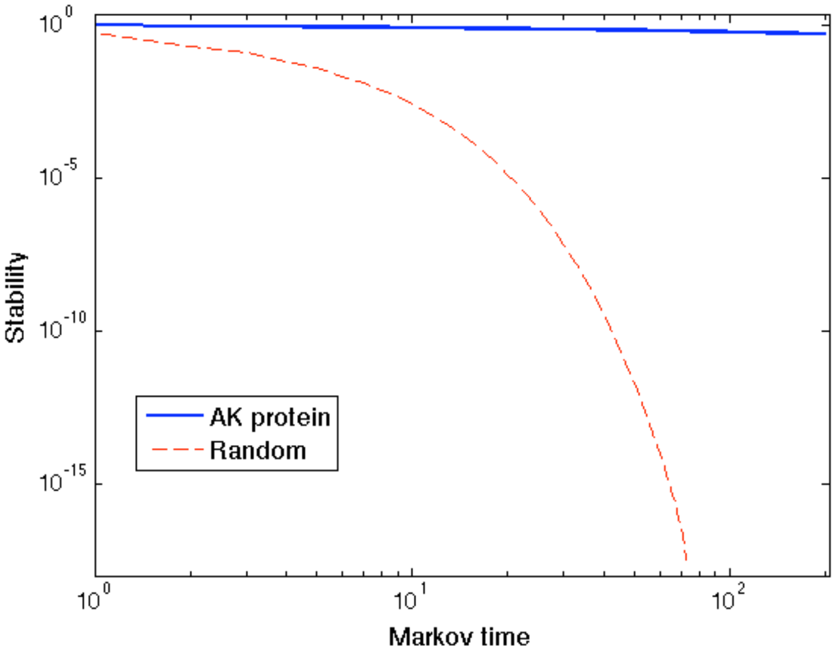


FIG. I: Stability curve of the graph an 2085-atom protein (AK) as shown in Figure 4 of the main text as compared with the stability curve of a random graph of the same size and similar average degree. We show that the stability for all partitions of the random graphs rapidly becomes negligible, indicating a clear lack of community structure in the random graph, while the stability of the partitions of the original protein graph decays slowly indicating that such partitions are indeed good communities.

2. **Graphs with particular community structure:**

Even though, as explained in the text (see also below), stability would typically tend towards a two-way (normalized Fiedler) partition as $t \to \infty$, this does *not* imply that this two-way partition will be necessarily identified as a relevant community structure. Indeed, in the case where such a partition is not relevant, it will have a very low value of the stability and thus it will be discounted as irrelevant in terms of community structure.

Example 1: An example of this behavior is given in Figure II below, where we consider a graph with no good two-way community due to its three-fold symmetry: a network formed by 3 complete graphs with 10 nodes (K_{10}) connected in a cycle with weights $1, 1+\epsilon, 1-\epsilon$. When $\epsilon \ll 1$, there is a clear three-community structure but barely any two-community structure. In this case, stability shows that that the time from which the two-community partition becomes dominant is very large and, in addition, the stability of the two-way partition is always below the epsilon-machine. Therefore, the three-community will be observed to be optimal at all times, while the two-community partition will not be observed in practice. If the graph is strictly symmetric ($\epsilon = 0$), then the two-community does not even appear as a potential community structure (this follows from the strict degeneracy of the eigenspectrum of the graph in the symmetric case). On the other hand, when $\epsilon \simeq 1$, there is a clear two-way community and this is shown by the stability curve, in which we observe a crossover from the three-way community to a two-way community at a finite time.

Example 2: The behavior observed in Figure II extends naturally to the case where the coarsest level of the community structure contains three or more subsets, rather than two. In the case when the graph has a dominant community structure with k communities at the coarsest level, its spectrum will have a group of $k-1$ equal (or very close) eigenvalues $\lambda_2, \ldots, \lambda_k$. In that case, the longest time scale of the stability curve, corresponding to two communities, will appear only at exceedingly long times, namely, for times t such that $\lambda_3^t \ll \lambda_2^t$. By the time this partition becomes dominant, the value of the stability will have dropped to values practically undistinguishable from zero, because λ_2^t will itself be negligible. There again, the two-way partition will not be considered as a good community following from the values of the stability curve.

We note that an example of this behavior was already presented in Fig. 3 of the main text: the Ravasz-Barabasi scale-free graph, which has an intrinsic five-fold symmetry. As one can see in Fig. 3, the stability curve unfolds the significant community structure (125 → 25 → 5) at different time scales. As long times, the five-fold

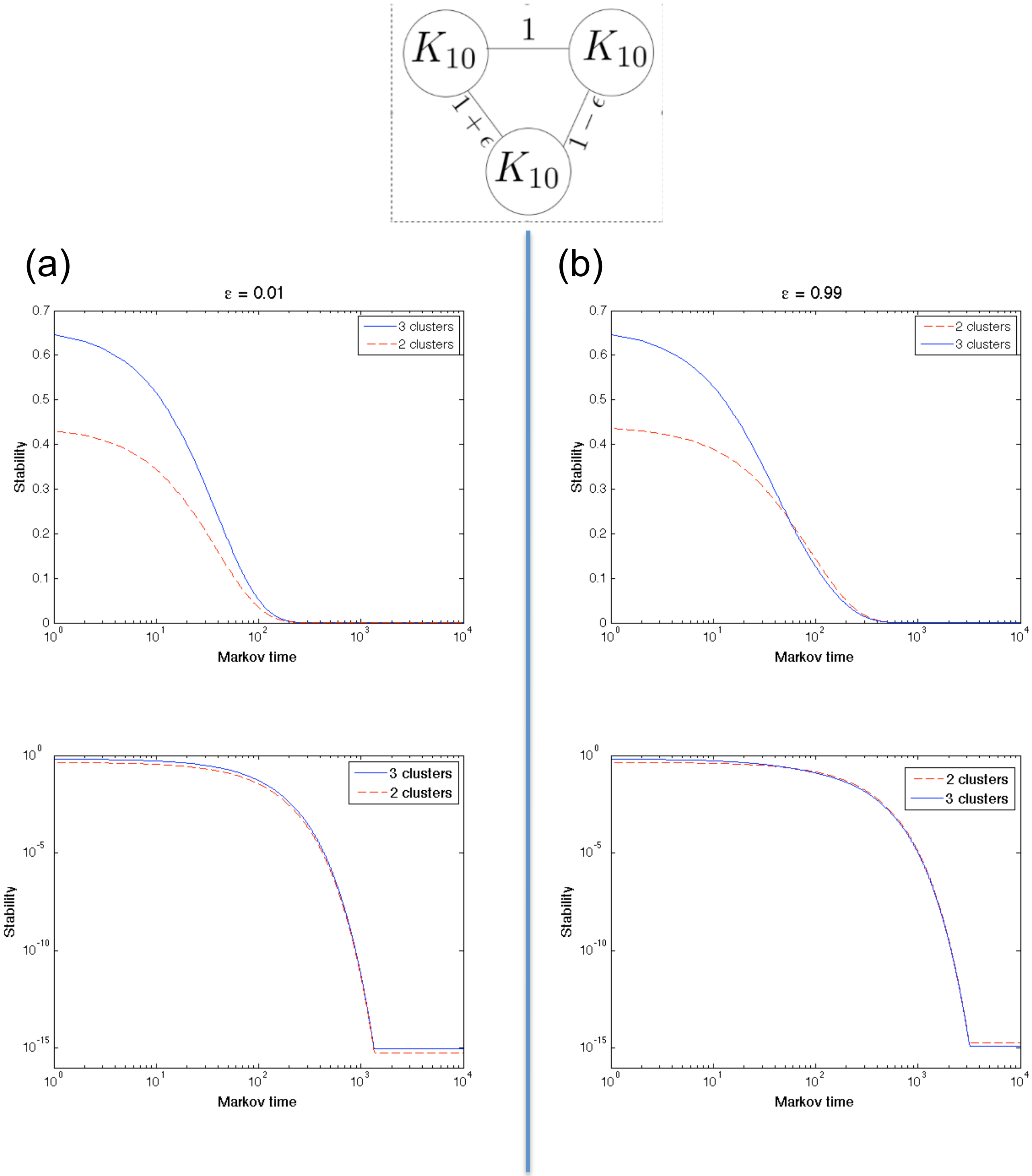


FIG. II: Stability curves for a graph formed by 3 complete graphs with 10 nodes (K_{10}) connected in a cycle with weights $1, 1+\epsilon, 1-\epsilon$ *(top panel)*. *(a)* For $\epsilon = 0.01$, the only relevant community structure is three-way. Indeed, stability shows that the three-way partition is dominant over a broad span of time scales, with significant values of its stability. Stability also shows that the time above which the two-community partition becomes dominant is so large that the stability of the two-way partition is always below the epsilon-machine. Therefore, the three-community will be observed to be optimal at all times, while the two-community partition will not be classed as a good community in practice at any time scale. The top and bottom figures represent the same data but the bottom subfigure is in log-log scale. *(b)* For $\epsilon = 0.99$, there is a clear two-way community and this is shown by the stability curve, in which we observe a crossover from the three-way community to a two-way community at a finite time with a non-negligible value of the stability. The bottom subfigure is again in log-log scale.

community is the relevant community: it dominates the stability curve until the stability reaches negligible values, indicating that the four-way, three-way and two-way partitions do not represent significant communities. This can be seen in more detail in log-log plot in Figure III, where we see the decay of the stability to the epsilon precision.

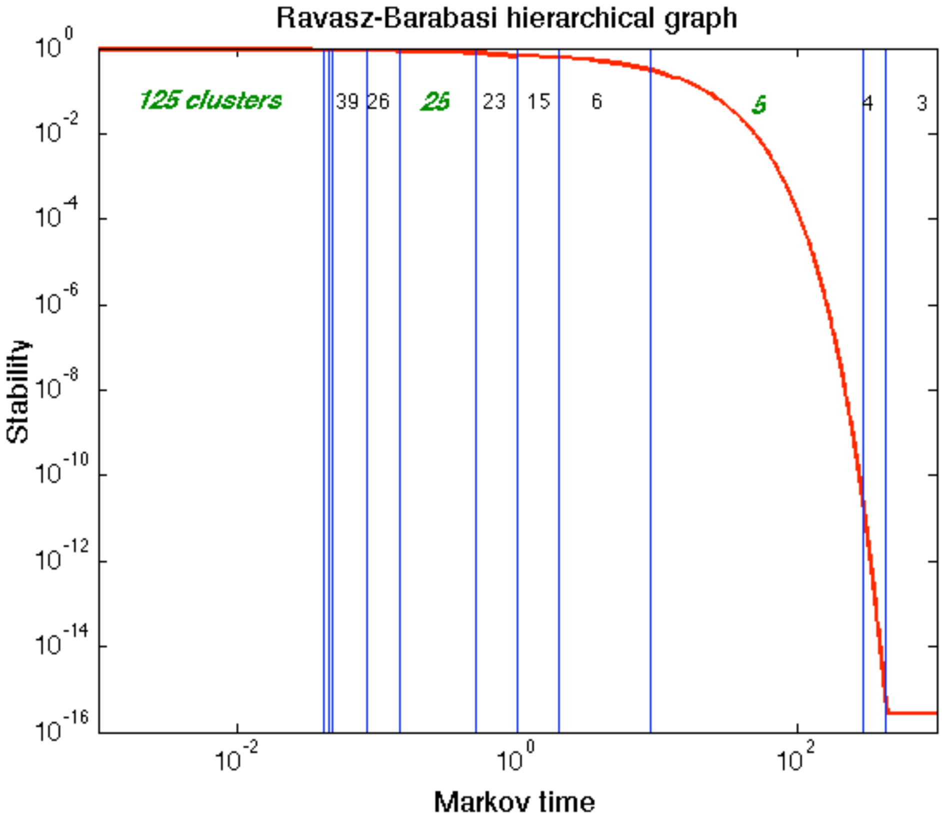


FIG. III: The stability curve of the hierarchical Ravazs-Barabasi network presented in Fig. 3 of the paper. This diagram is the same as in Fig 3 of the main paper, only in log-log scale.

III. COMPUTATIONAL ASPECTS OF THE OPTIMIZATION OF STABILITY

A. NP-hardness and heuristics for stability maximization

Most of the non-trivial problems in clustering and partitioning, including modularity optimization, are NP-hard. Similarly, the optimization of stability over the space of possible partitions is also an NP-hard problem. This means that stability optimization has to be implemented algorithmically through heuristics, without global guarantees of optimality, except for simple asymptotic values such as the stability at $t = 0$ or as $t \to \infty$ discussed above.

The computational complexity of the global optimization problem has only fueled further research into these problems leading to the proposal, over the last twenty years, of efficient algorithmic heuristics for clustering and, more recently, for modularity optimization. In this respect, one of the key advantages of stability is that it can be used to rank the partitions obtained through the use of different algorithms and, if advantageous, one can pick different clustering algorithms at different timescales to obtain the most relevant communities.

In addition, it is important to mention that stability can be shown to be equivalent to the modularity of a time-dependent graph [6]. This means that any of the recently proposed heuristic methods used for the modularity optimization can be used to maximize the stability without further modification.

B. Stability of partitions obtained through different heuristic algorithms

Some of the issues above have been explored in the main text. Figure 2 of the main text presents the use of stability to evaluate the partitions that are obtained by a variety of well-known algorithms that find ‘good’ clusterings according to different heuristics. Stability provides a way to rank the partitions so obtained and, conversely, different algorithms (and heuristics) can be used to find partitions that optimize stability at different scales.

Brief description of the methods used in Figure 2 of the main text:

- **Shi-Malik:** The first method is a variant of Shi-Malik’s algorithm [8], which is a recursive spectral method where the graph is divided in two parts from the normalized Fiedler vector. The entries of the vector are sorted and a linear search is performed to find the cutting point that maximises the Normalized Cut of the resulting two-way partition. Then the resulting cluster with the highest normalized second eigenvalue of the adjacency matrix is split again until we get k clusters.

- **KVV:** The second method is a variant of an algorithm introduced by Kannan, Vempala and Vetta [13], and therefore called the KVV method. It proceeds similarly to Shi-Malik’s algorithm, by dividing every cluster in two parts until we obtain k clusters. It aims to optimize the conductance for the partitions. The conductance of a two-way clustering is the number of edges between the clusters divided by the minimum sum of degrees of nodes in a cluster. A cluster is split by first ordering the vertices according to the eigenvector as in the first method, then by choosing the cutting point that optimizes conductance. At every step, the split that results in the smallest conductance is chosen.

  Note: Our implementation of the Shi-Malik and KVV algorithms is equivalent to those proposed in the SpectraLIB toolbox for Matlab, written by Verma and Maravina [14].

- **Newman (without Kernighan-Lin):** The third method is essentially Newman’s spectral method to optimize modularity [15], which relies on the dominant eigenvector of the matrix $\Pi M - \pi^T \pi$. The algorithm proceeds recursively by finding bipartitions from the full graph in the way described in [15], starting by the splitting that leads to the largest increase of modularity, until we have the required number of clusters or no division leads to an increase of modularity. To keep things simple, we do not fine-tune the result with the Kernighan-Lin-like procedure.

- **Newman-Girvan:** The last method is the Newman-Girvan algorithm [4] that deletes the edge with highest betweenness centrality until k clusters are obtained.

Another algorithm (see Figure IV) — The Louvain method: We also note that we have used other algorithms, including agglomerative heuristics such as those employed by the Louvain method [6] and we have obtained consistent results. Therefore, different heuristics can be used for different graphs and, indeed, for different timescales to select the best partitions according to their stability.

C. Spectral clustering as a heuristic for the optimization of Stability

It is interesting to note that two of the heuristics used above (Shi-Malik [8] and Newman [15]) are based on successive spectral bi-partitioning and are used to obtain partitions that can then be evaluated at different timescales. In this sense, it can be said that hierarchical spectral clustering (in the two flavors given by Shi-Malik and Newman) is being used here as a heuristic to optimize stability at different times and, as such heuristic, it performs well in many cases. We can provide further insight into the use of this heuristic by considering its connection with previous research by Newman and Shi and Malik:

- **Spectral clustering for modularity optimization:** Newman [11] observed that finding a two-way partition with optimal modularity can be written (using our notation) as the following optimization problem:

  $$\max v^T(\Pi M - \pi^T \pi)v \text{ subject to } v_i = \pm 1, \ \forall i.$$

  Once an optimal v is found, the nodes are partitioned according to the sign of the entries of v. This combinatorial optimization problem is NP-hard but relaxing the constraint $v_i = \pm 1$ to the weaker constraint $v^T v = n$ leads to an easy optimisation problem, whose solution is precisely the dominant eigenvector of $\Pi M - \pi^T \pi$. Grouping the nodes according to the sign of this eigenvector, together with a clever recursion, allowed Newman to devise a spectral heuristics to optimize modularity.

- **Spectral methods related to Normalized Cut:** Similarly, we observe that the constraint $v_i = \pm 1$ can also be relaxed to $v^T \Pi v = 1$, which now leads to an optimization problem whose solution is the partition that groups the nodes according to the dominant eigenvector of $M - \mathbf{1}\pi$, where $\mathbf{1}$ is the column vector of ones. This is precisely the second eigenvector of M, i.e., the normalized Fiedler vector. Similarly to Newman's heuristic, one can then use the sign of the *normalized* Fiedler vector to create a recursive spectral method that will serve as a heuristic to optimize modularity. This also shows that this heuristic, used by Shi and Malik [8] to optimize NCut, is also a reasonable heuristic for finding the modularity-optimal two-way partition.

- **Spectral methods for Stability optimization:** As stated above, it is easy to see that the stability of a graph at time t can be expressed as the modularity of a graph with time-dependent weights. This can be seen by considering that R_t for a graph of adjacency matrix $\Pi M = A/2m$ is the modularity of a graph whose adjacency matrix is ΠM^t. Therefore, we can use both Newman's and Shi-Malik's spectral algorithms as reasonable heuristics to optimise stability at all times t, since R_t can always be interpreted as a 'modularity'. Obviously, according to the time t, it will be necessary to proceed with more or fewer subdivisions. That is, it will be likely that partitions obtained by many successive bi-partitions according to normalized Fiedler (Shi-Malik) will have high stability at short times. At large times, the partitions with higher stability will be those obtained after a few recursive bi-partitions. And, in particular, as $t \to \infty$ we have seen above that the subdivision into only two sets according to the normalized Fiedler vector is typically optimal, although as explained above, its stability might be so low as to make it irrelevant as a community.

The above discussion gives insight and justifies the use of different variants of Shi-Malik's algorithm and Newman's spectral algorithm (without the Kernighan-Lin step) as heuristics for stability optimization in the text.

D. Robustness of Stability and the community structure obtained with different computational methods

Since stability optimization is NP-hard, all the algorithms used in this paper are heuristics. As such, it is important to check that the value of stability obtained through the use of different algorithms is robust. This is shown in Figure 2 of the main paper, where we compare algorithms based on different heuristics, yet the overall values of the stability (and its time-dependence) are broadly similar. The fact that a variety of methods/heuristics give comparable estimates of the optimal stability gives further support to its use to establish the relevance of different communities. Future work will be devoted to the characterization of the robustness of stability at different time scales.

In addition to the consistency of the estimates of stability, it is also important to check that the relevant communities are robust and do not depend much on the particular algorithm used. It is indeed the case that relevant partitions are robust in this sense (as well as in a more general sense) and this property can be used as a criterion for the selection of relevant communities. This exploration falls beyond the scope of this paper and will be studied in future work. However, we show here the illustrative example of Fig. 1 of the paper, and compute the community structure obtained by two very different heuristics: the KVV method, a divisive method that splits the graph recursively based on the minimization of conductance between the subgraphs of the partition, and the multiscale Louvain algorithm, a greedy agglomerative method that starts from isolated nodes and aggregates them in a locally optimal fashion [6].

In Figure IV, we show that the stability curve obtained using both methods has similar values and a similar time dependence. We also compare the partitions obtained in each case, using the relative information distance, and show that the relevant partitions of similar size are broadly robust even when obtained with differing algorithmic heuristics [6].

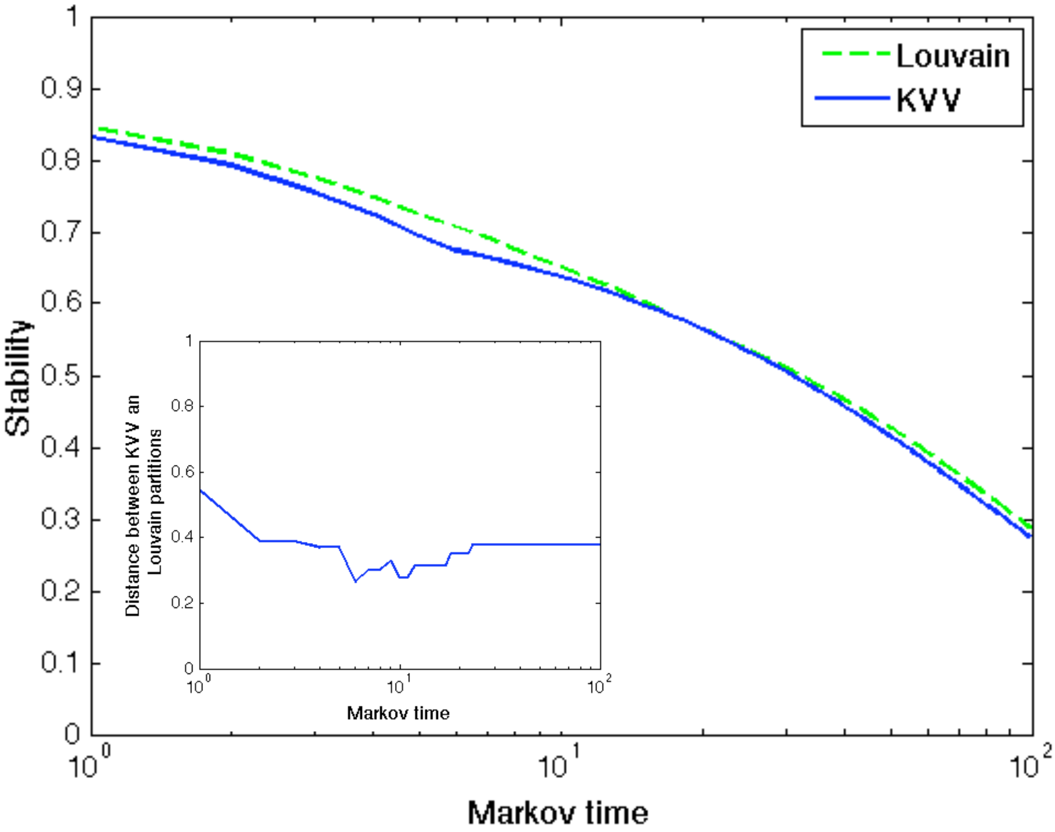


FIG. IV: The stability curves of the collaboration network (Fig. 1 of the main text) obtained with two algorithms that use very different optimization heuristics: a global divisive method (KVV), as in the main text, and a greedy agglomerative algorithm (Louvain). The fact that the value and temporal dependence of the stability are very similar in both cases indicates that our measure is robust to the optimization algorithm employed. We have used a variety of methods (see list above) and the results are similar (see Fig. 2 of the main text). *Inset:* The difference between the partitions obtained by the KVV method and by the Louvain method is small, as measured by the relative information distance. The relative information distance of two partitions P and Q is given by $1 - I(P;Q)/H(P,Q)$, where $I(P;Q)$ is the mutual information of the partitions, and $I(P;Q)$ is their joint entropy. It is normalized and takes values between 0 and 1. Low values of this distance indicate that the partitions are similar.

IV. TECHNICAL COMMENTS ON SOME OF THE DATA PRESENTED IN THE EXAMPLES AND FIGURES OF THE MAIN TEXT

A. Obtaining the atomistic graph of the protein AK used in Example 3 (Figure 4 of the main text)

The protein Adenylate Kinase (AK), with 2085 atoms, was used in Example 3. The nodes of the graph correspond to each of the atoms and the edges connecting them were obtained as follows [16]. Firstly, the coordinate file for the open conformation (PDB 4ake) was downloaded from the Protein Databank [17] and hydrogens were added using the standard GROMACS hydrogen placement based on the pdb2gmx function [18] which has been shown to perform well [19]. Then the protein was equilibrated using molecular dynamics with the force-field Gromos96-43a1 in solvent with the SCP water model. Finally, the edges are distance constraints detected using the software FIRST (Floppy Inclusions and Rigid Sub-structure Topography) [20], version 5.2. FIRST identifies constraints based on covalent, hydrogen bonds, salt bridges and hydrophobic contacts in the body-bar formalism. The exact mapping of the body-bar formalism to a bond-bending network [21] was subsequently used to extract the distance constraints and, where appropriate, assign angle constraints. The standard energy cutoff of -1 kcal/mol was used.

B. Details of the network science collaboration graph used in Example 1 (Figure 1 of the main text)

The graph of collaborations is from Newman [15]. The names of the researchers in the 21 groups in Figure 1*A* are given in the appendix below grouped also according to the colors of the dendrogram presented in Figure 1*B*. Groups that merge quickly (e.g., clusters 15 and 18) are surrounded by darker boxes. The nodes are ordered by decreasing degree (shown in the tables) within each cluster. For interested readers who are also researchers in the field of network theory, it is interesting to consider these coarser meta-communities in the light of research sub-areas and also in terms of closeness of methodological approaches.

[1] E. H. Simpson. Measurement of diversity. *Nature*, 163:688, 1949.

[2] A. O. Hirschman. The Paternity of an Index. *The American Economic Review*, 54(5):761, 1964.

[3] A. Rényi. On Measures of Entropy and Information. *Proc. Fourth Berkeley Symp. Math. Stat. and Probability*, Vol. 1. Berkeley, CA: University of California Press, pp. 547-561, 1961.

[4] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2), 2004.

[5] J. Reichardt and S. Bornholdt. Statistical mechanics of community detection. *Phys. Rev. E*, 74:016110, 2006.

[6] R. Lambiotte, J.-C. Delvenne and M. Barahona. Laplacian Dynamics and Multiscale Modular Structure in Networks. *arXiv:0812.1770*, 2008.

[7] A. Arenas, A. Fernndez, and S. Gómez. Multiple resolution of the modular structure of complex networks. *New J. Phys.*, 10:05039, 2008.

[8] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[9] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(98):298–305, 1973.

[10] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(100):619–633, 1975.

[11] M. E. J. Newman. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA*, 103:8577–8582, 2006.

[12] S. Fortunato and C. Castellano. *Encyclopedia of Complexity and System Science*, chapter Community structure in graphs. Springer, 2008.

[13] R. Kannan, S. Vempala, and A. Vetta. On clusterings : Good, bad and spectral. In *Proceedings of the Symposium on Foundations of Computer Science*, 2000.

[14] D. Verma and T. Maravina. Spectralib - package for symmetric spectral clustering.

[15] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 74(3), 2006.

[16] J. Costa. Ph.D. Thesis, Imperial College London (2008).

[17] H. Berman *et al.* The Protein Data Bank. *Nucleic Acids Res.*, 28, 235 (2000).

[18] D.V.D. Spoel *et al.* GROMACS: fast, flexible, and free. *J. Comp. Chem.*, 26, (2005).

[19] L.R. Forrest and B. Honig. An assessment of the accuracy of methods for predicting hydrogen positions in protein structures. *Proteins*, 61, 296 (2005).

[20] D.J. Jacobs, A.J. Rader, L.A. Kuhn and M.F. Thorpe. Protein flexibility predictions using graph theory. *Proteins*, 44, 150 (2001).

[21] W. Whiteley. Counting out to the flexibility of molecules. *Phys. Biol.*, 2, S116 (2005).

	Cluster 15:
8	KRAPIVSKY, P
7	REDNER, S
5	RODGERS, G
4	TADIC, B
2	THURNER, S
1	ANTAL, T
1	ERGUN, G
1	LEYVRAZ, F
1	DARBYDOWMAN, K
1	VAZQUEZ, F

	Cluster 18:
3	TOROCZKAI, Z
3	BASSLER, K
2	PACZUSKI, M
2	BENNAIM, E
1	BAIESI, M
1	FRAUENFELDER, H
1	CORRAL, A
1	KOZMA, B
1	HENGARTNER, N
1	KORNISS, G

	Cluster 16:
11	LATORA, V
6	CRUCITTI, P
6	MARCHIORI, M
3	FORTUNA, L
3	LAROSA, M
3	RAPISARDA, A
2	FRASCA, M
2	PORTA, S
1	BUCOLO, M
1	CARUSO, F
1	COSENZA, S
1	STAGNI, C
1	USAI, L
1	SPATA, A
1	PLUCHINO, A
1	TIERI, P
1	VALENSIN, S
1	CASTELLANI, G
1	REMONDINI, D
1	FRANCESCHI, C

	Cluster 4:
7	HAVLIN, S
6	COHEN, R
5	BENAVRAHAM, D
1	EREZ, K
1	ROZENFELD, A
1	SCHWARTZ, N
1	SONG, C
1	MAKSE, H

	Cluster 17:
30	BARABASI, A
18	JEONG, H
11	KAHNG, B
10	OLTVAI, Z
9	VICSEK, T
8	ALBERT, R
6	RAVASZ, E
6	GOH, K
6	KIM, D
5	ALMAAS, E
4	NEDA, Z
4	HONG, H
3	KULKARNI, R
3	BIANCONI, G
3	SCHUBERT, A
3	DEMENEZES, M
3	DUNNE, J
3	WILLIAMS, R
3	FARKAS, I
3	OH, E
3	YOOK, S
2	STROUD, D
2	BERLOW, E
2	MOUKARZEL, C
2	MARTINEZ, N
2	DERENYI, I
2	TOMBOR, B
2	KIM, J
2	RHO, K
2	BENJACOB, E
1	ALBERT, I
1	NAKARADO, G
1	KOVACS, B
1	PENNA, T
1	DEZSO, Z
1	DOBRIN, R
1	BEG, Q
1	GHIM, C
1	LEE, D
1	GORMAN, S
1	MASON, S
1	JUNG, S
1	KIM, S
1	PARK, Y
1	KINNEY, R
1	MACDONALD, P
1	MARODI, M
1	DOVIO, F
1	PODANI, J
1	SZATHMARY, E
1	SOMERA, A
1	MONGRU, D
1	SHEFI, O
1	GOLDING, I
1	SEGEV, R
1	AYALI, A
1	CZIROK, A
1	COHEN, I
1	SHOCHET, O
1	WUCHTY, S
1	TU, Y

	Cluster 1:
9	HOLME, P
8	KIM, B
7	SNEPPEN, K
4	TRUSINA, A
4	CHOI, M
3	MASLOV, S
3	MINNHAGEN, P
2	GHOSHAL, G
2	HUSS, M
2	YOON, C
2	HAN, S
2	ROSVALL, M
1	BAK, P
1	ERIKSEN, K
1	SIMONSEN, I
1	PARK, H
1	CHUNG, J
1	KUMAR, S
1	ZALIZNYAK, A
1	LEICHT, E

	Cluster 10:
4	RAGHAVAN, P
4	RAJAGOPALAN, S
4	TOMKINS, A
3	KUMAR, R
1	BRODER, A
1	MAGHOUL, F
1	STATA, R
1	WIENER, J
1	SIVAKUMAR, D
1	UPFAL, E

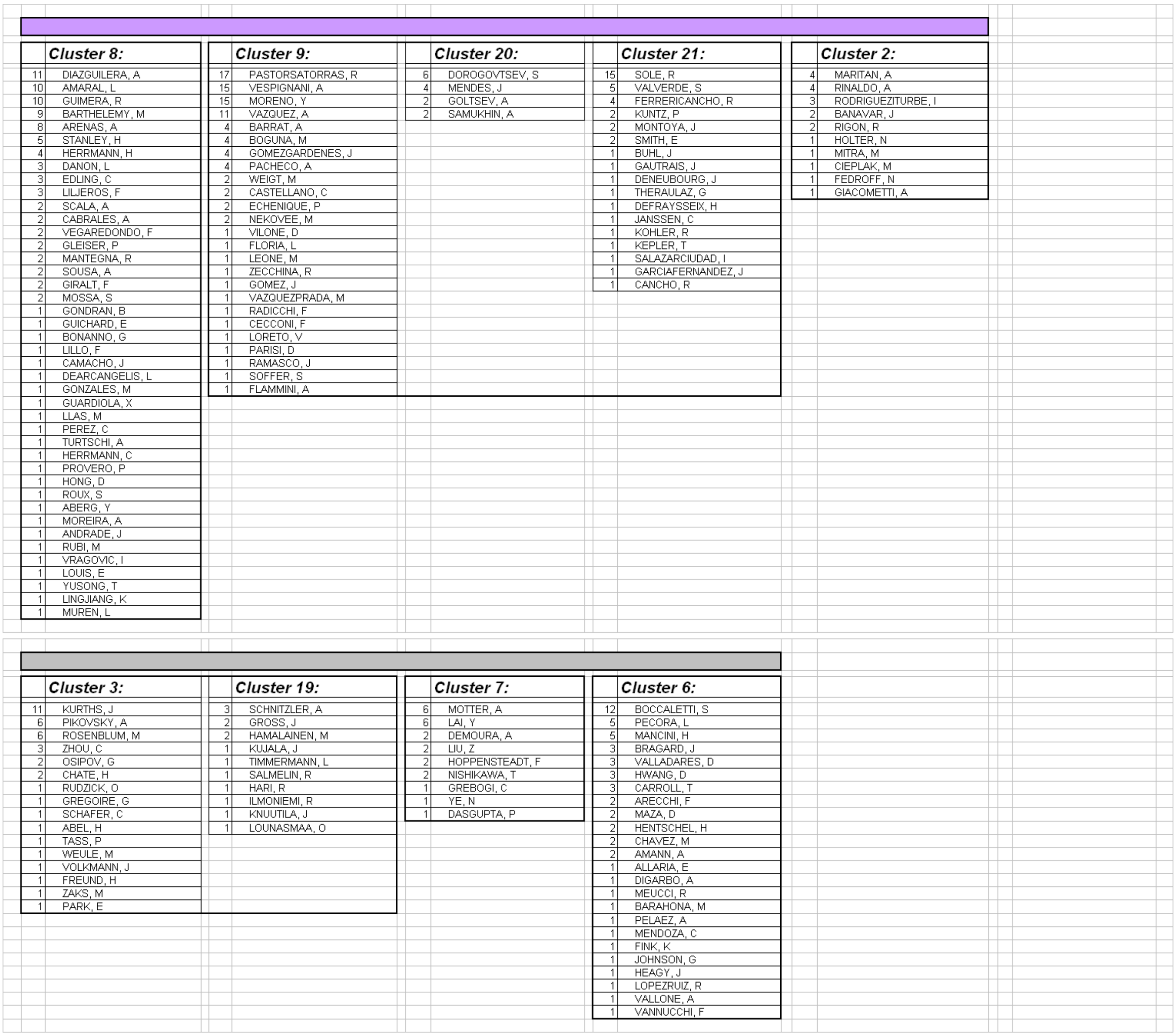
	Cluster 12:
4	LAWRENCE, S
3	KLEINBERG, J
3	GILES, C
2	FLAKE, G
1	COETZEE, F
1	PENNOCK, D
1	GLOVER, E

	Cluster 13:
23	NEWMAN, M
9	WATTS, D
8	STROGATZ, S
5	MOORE, C
4	DODDS, P
3	GIRVAN, M
2	CALLAWAY, D
2	MIROLLO, R
1	ANCELMEYERS, L
1	MARTIN, M
1	SCHRAG, S
1	HOPCROFT, J
1	CLAUSET, A
1	MUHAMAD, R
1	ROTHMAN, D
1	SABEL, C
1	GASTNER, M
1	JIN, E
1	KALAPALA, V
1	SANWALANI, V
1	LUSSEAU, D
1	MATTHEWS, P
1	FORREST, S
1	BALTHROP, J
1	PARK, J
1	PORTER, M
1	MUCHA, P
1	WARMBRAND, C
1	YEUNG, M

	Cluster 14:
9	STAUFFER, D
3	HOLYST, J
3	DAFONTOURACOSTA, L
2	BERNARDES, A
2	FORTUNATO, S
2	FRONCZAK, A
2	AHARONY, A
1	ALEKSIEJUK, A
1	COSTA, U
1	ARAUJO, A
1	DIAMBRA, L
1	FRONCZAK, P
1	JEDYNAK, M
1	SIENKIEWICZ, J
1	ADLER, J
1	MEYERORTMANNS, H

	Cluster 11:
5	KERTESZ, J
3	KASKI, K
2	ONNELA, J
1	LAHTINEN, J
1	CHAKRABORTI, A
1	KANTO, A
1	JARISARAMAKI, J
1	SZABO, G
1	ALAVA, M

	Cluster 5:
7	CALDARELLI, G
4	CAPOCCI, A
4	DELOSRIOS, P
3	MUNOZ, M
3	GARLASCHELLI, D
2	BATTISTON, S
2	SERVEDIO, V
2	PIETRONERO, L
2	MARRO, J
1	CATANZARO, M
1	COCCETTI, F
1	COLAIORI, F
1	DELUCIA, M
1	BOTTACCIO, M
1	MONTUORI, M
1	DONETTI, L
1	CASTRI, M
1	LOFFREDO, M
1	DICKMAN, R
1	PETERMANNN, T
1	TORRES, J
1	GARRO, P



	Cluster 8:
11	DIAZGUILERA, A
10	AMARAL, L
10	GUIMERA, R
9	BARTHELEMY, M
8	ARENAS, A
5	STANLEY, H
4	HERRMANN, H
3	DANON, L
3	EDLING, C
3	LILJEROS, F
2	SCALA, A
2	CABRALES, A
2	VEGAREDONDO, F
2	GLEISER, P
2	MANTEGNA, R
2	SOUSA, A
2	GIRALT, F
2	MOSSA, S
1	GONDRAN, B
1	GUICHARD, E
1	BONANNO, G
1	LILLO, F
1	CAMACHO, J
1	DEARCANGELIS, L
1	GONZALES, M
1	GUARDIOLA, X
1	LLAS, M
1	PEREZ, C
1	TURTSCHI, A
1	HERRMANN, C
1	PROVERO, P
1	HONG, D
1	ROUX, S
1	ABERG, Y
1	MOREIRA, A
1	ANDRADE, J
1	RUBI, M
1	VRAGOVIC, I
1	LOUIS, E
1	YUSONG, T
1	LINGJIANG, K
1	MUREN, L

	Cluster 9:
17	PASTORSATORRAS, R
15	VESPIGNANI, A
15	MORENO, Y
11	VAZQUEZ, A
4	BARRAT, A
4	BOGUNA, M
4	GOMEZGARDENES, J
4	PACHECO, A
2	WEIGT, M
2	CASTELLANO, C
2	ECHENIQUE, P
2	NEKOVEE, M
1	VILONE, D
1	FLORIA, L
1	LEONE, M
1	ZECCHINA, R
1	GOMEZ, J
1	VAZQUEZPRADA, M
1	RADICCHI, F
1	CECCONI, F
1	LORETO, V
1	PARISI, D
1	RAMASCO, J
1	SOFFER, S
1	FLAMMINI, A

	Cluster 20:
6	DOROGOVTSEV, S
4	MENDES, J
2	GOLTSEV, A
2	SAMUKHIN, A

	Cluster 21:
15	SOLE, R
5	VALVERDE, S
4	FERRERICANCHO, R
2	KUNTZ, P
2	MONTOYA, J
2	SMITH, E
1	BUHL, J
1	GAUTRAIS, J
1	DENEUBOURG, J
1	THERAULAZ, G
1	DEFRAYSSEIX, H
1	JANSSEN, C
1	KOHLER, R
1	KEPLER, T
1	SALAZARCIUDAD, I
1	GARCIAFERNANDEZ, J
1	CANCHO, R

	Cluster 2:
4	MARITAN, A
4	RINALDO, A
3	RODRIGUEZITURBE, I
2	BANAVAR, J
2	RIGON, R
1	HOLTER, N
1	MITRA, M
1	CIEPLAK, M
1	FEDROFF, N
1	GIACOMETTI, A

	Cluster 3:
11	KURTHS, J
6	PIKOVSKY, A
6	ROSENBLUM, M
3	ZHOU, C
2	OSIPOV, G
2	CHATE, H
1	RUDZICK, O
1	GREGOIRE, G
1	SCHAFER, C
1	ABEL, H
1	TASS, P
1	WEULE, M
1	VOLKMANN, J
1	FREUND, H
1	ZAKS, M
1	PARK, E

	Cluster 19:
3	SCHNITZLER, A
2	GROSS, J
2	HAMALAINEN, M
1	KUJALA, J
1	TIMMERMANN, L
1	SALMELIN, R
1	HARI, R
1	ILMONIEMI, R
1	KNUUTILA, J
1	LOUNASMAA, O

	Cluster 7:
6	MOTTER, A
6	LAI, Y
2	DEMOURA, A
2	LIU, Z
2	HOPPENSTEADT, F
2	NISHIKAWA, T
1	GREBOGI, C
1	YE, N
1	DASGUPTA, P

	Cluster 6:
12	BOCCALETTI, S
5	PECORA, L
5	MANCINI, H
3	BRAGARD, J
3	VALLADARES, D
3	HWANG, D
3	CARROLL, T
2	ARECCHI, F
2	MAZA, D
2	HENTSCHEL, H
2	CHAVEZ, M
2	AMANN, A
1	ALLARIA, E
1	DIGARBO, A
1	MEUCCI, R
1	BARAHONA, M
1	PELAEZ, A
1	MENDOZA, C
1	FINK, K
1	JOHNSON, G
1	HEAGY, J
1	LOPEZRUIZ, R
1	VALLONE, A
1	VANNUCCHI, F